Galaxy evolution modelling with simulated deep fields, Bayesian inference and dimensional reduction through neural network

Florian Livet

Institut d'Astrophysique de Paris

03/04/2019

PhD advisors : Valérie de Lapparent and Damien Le Borgne

What is a galaxy?

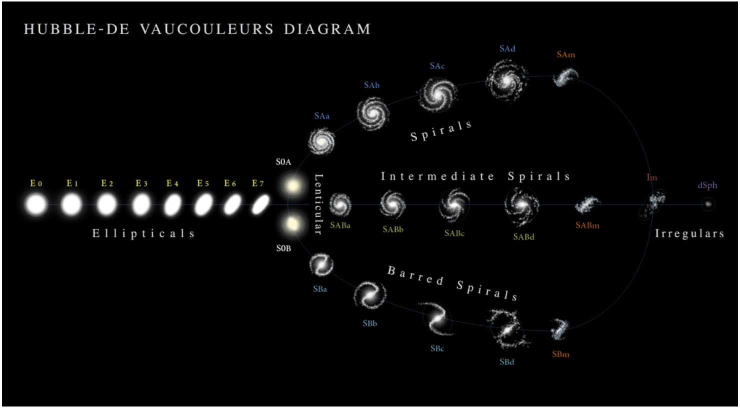
- Stars (between 10^8 to 10^{11})
- Stellar remnants
- Interstellar gas
- Dust
- Black holes (stellar and super massive)
- Dark matter?

$\longrightarrow$ **GRAVITATION**

NGC 4414

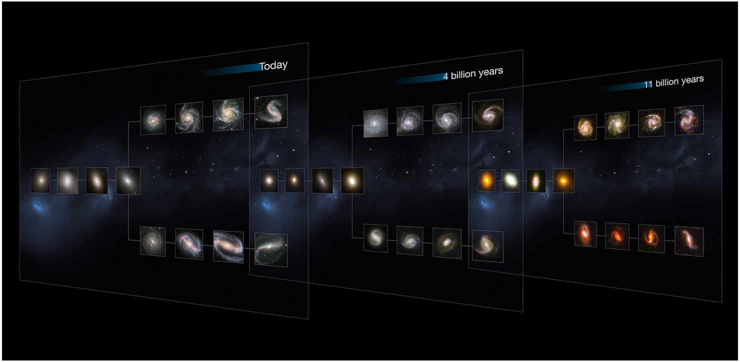
Morphology

- Galaxies are categorized according to their visual morphology ($-6 \rightarrow 11$).
- Bulge + Disk



Formation and evolution

- Hypothesis: Galaxy formation occured from tiny quantum fluctuations after the Big Bang.
- Clustering and merging = mass accumulation ⇒ determining their shape and structure.



Redshift

Because light cannot travel faster than $c = 3 \times 10^8 ms^{-1}$ then the distance depends on the time:

$$d(t) = a(t)d_0 \tag{1}$$

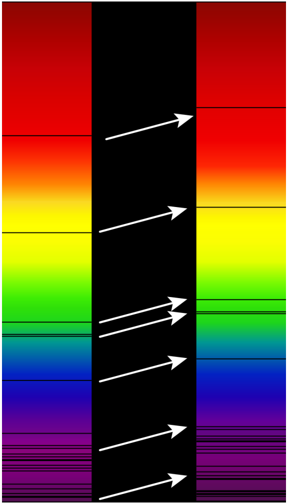
where

- t is the look-back time.
- d_0 is the distance at the reference time $t = 13.799 Gyr$.
- $a(t)$ is the scale factor at time t.

The redshift z of an object is then defined by:

$$a(t) = \frac{1}{1+z} \tag{2}$$

Redshift



Absorption lines in the visible spectrum of a supercluster of distant galaxies (right), as compared to absorption lines in the visible spectrum of the Sun (left).

Galaxy evolution modelling

SCHECHTER LUMINOSITY FUNCTION

Gives the number of galaxies per volume (Mpc^3) per magnitude.
$1\mathrm{MPC} = 3.086 \times 10^{22}$ m.

magnitude = measure of the luminosity of a celestial object.

$$\Phi(M)dM = 0.4\log(10)\phi^*(z)\exp\left[0.4\log(10)(M^*(z)-M)(1+\alpha) - \exp(0.4\log(10)(M^*(z)-M))\right]$$

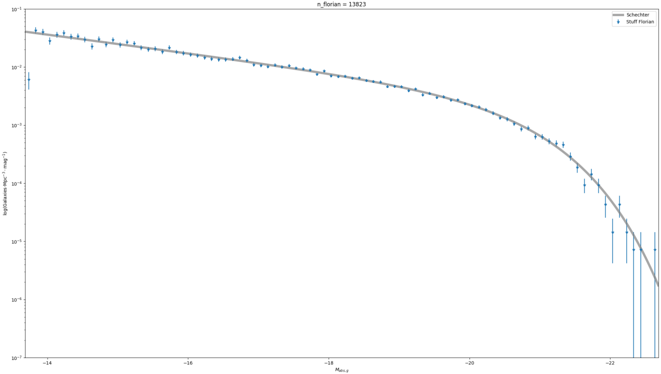
where

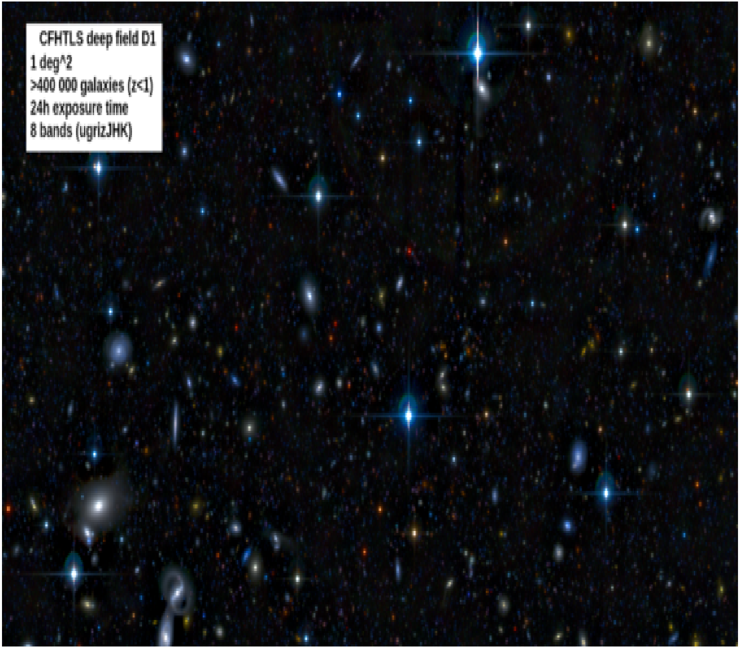
- z is the redshift
- $M^*(z) = M^* + M^*_{\mathrm{evol}}\log(1+z)$
- $\phi^*(z) = \phi^* \times (1+z)^{\phi^*_{\mathrm{evol}}}$

→ 5 free parameters: α, M^*, ϕ^*, M^*_{evol} and ϕ^*_{evol} **for each population** of galaxies.

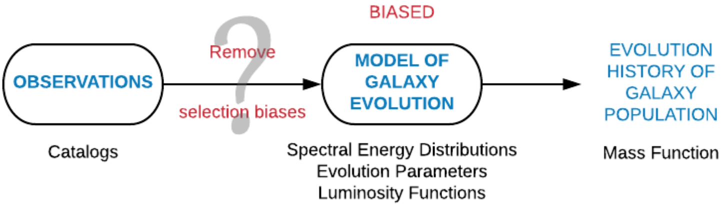
Galaxy counts

$\alpha = -1.4$, $M^* = -19.84$, $\phi^* = 0.1$, $M^*_{\text{evol}} = 0$ and $\phi^*_{\text{evol}} = 0$ for spiral galaxies.




CFHTLS deep field D1
1 deg^2
>400 000 galaxies (z<1)
24h exposure time
8 bands (ugrizJHK)

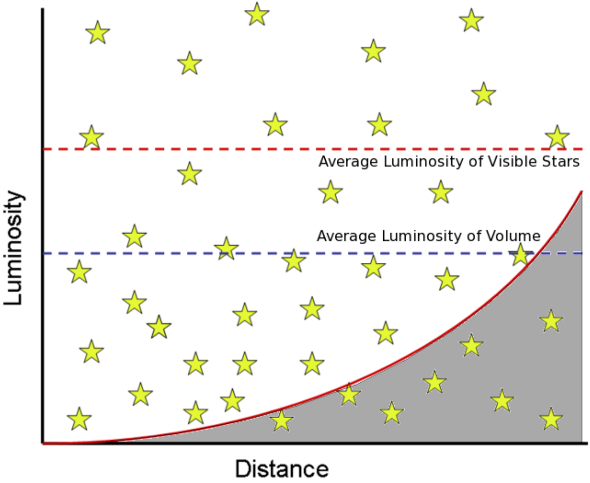
Classic approach: Inverse modelling



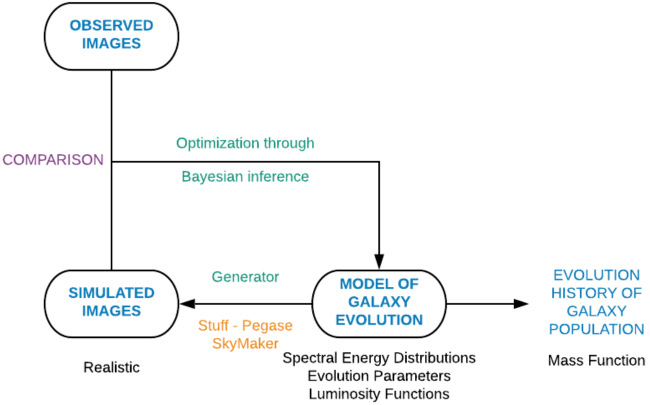
⚠ **SELECTION BIASES** ⚠

- Cosmological dimming $(1+z)^{-4}$
- Malmquist bias
- Eddington bias
- Expansion of the Universe (k-correction)
- Dust extinction
- Source occultation and confusion

Malmquist bias



Our approach: Forward modelling

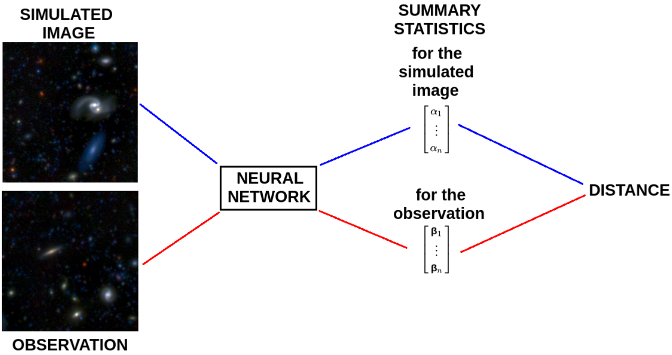


ROBUSTNESS

Simulated and observed images
have the same selection biases.

Comparison of simulated and observed images

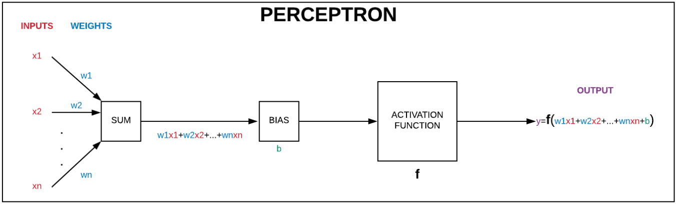
Charnock et al. 2018: "Automatic physical inference with information maximising neural networks".



$$\text{distance} = \sqrt{(\alpha-\beta)^T F(\alpha-\beta)} \tag{3}$$

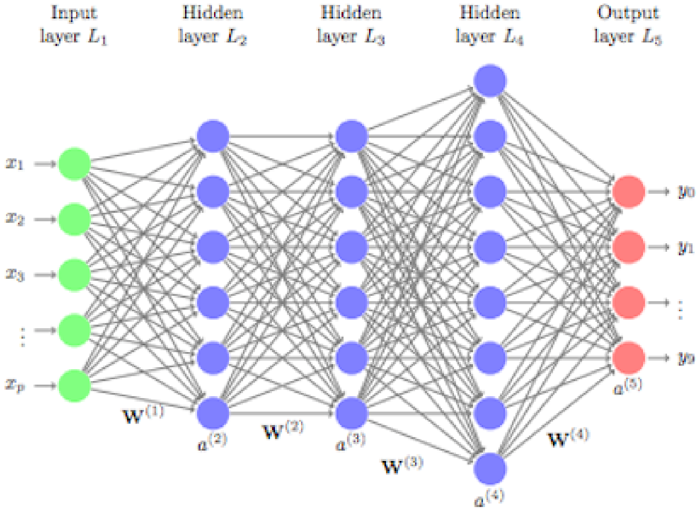
where F is the Fisher information matrix obtained by the trained network.

What is a neural network?

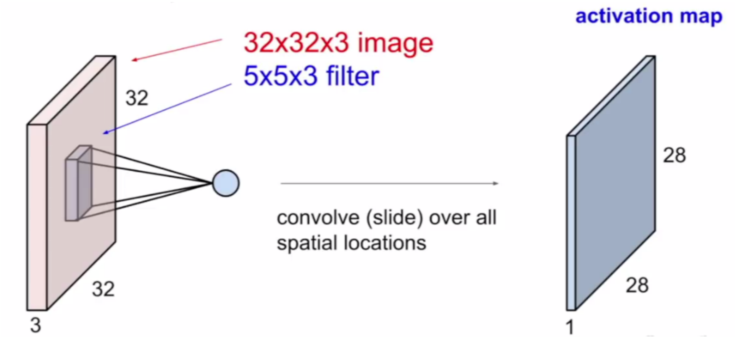


$$y = f\left(\sum_{i=1}^{n} w_i x_i + b\right) \tag{4}$$

Fully connected layers



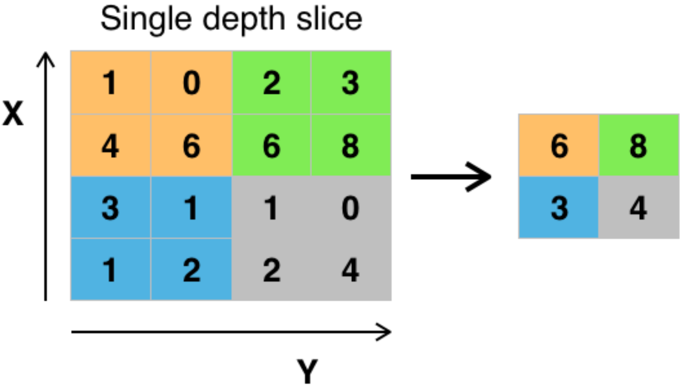
Convolutional layer



ADVANTAGE

- Study different patterns of an image.

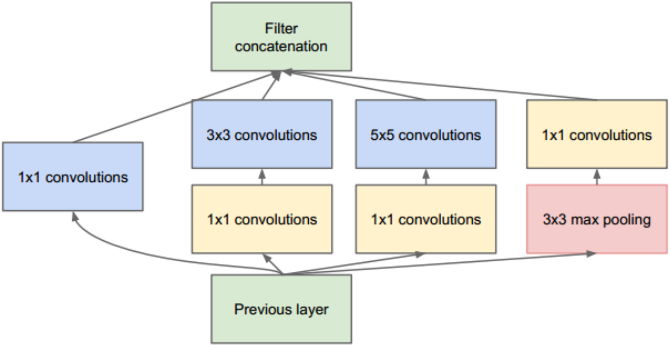
Pooling layer



ADVANTAGE

- Decrease spatial resolution of an image.

Inception module



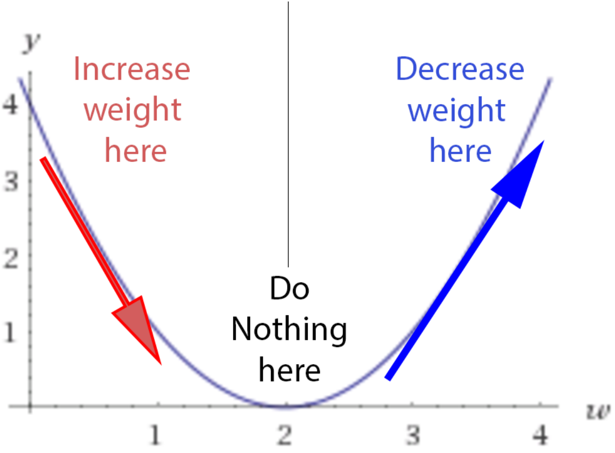
ADVANTAGE

- Study different patterns of an image for multiple spatial resolutions.

Inception network = several inception modules followed by several fully connected layers (deep network).

Backpropagation

- Loss function = calculates the correctness of the network outputs.
- Backpropagation = backward propagation of errors.
- Weights and biases update.



Fisher information

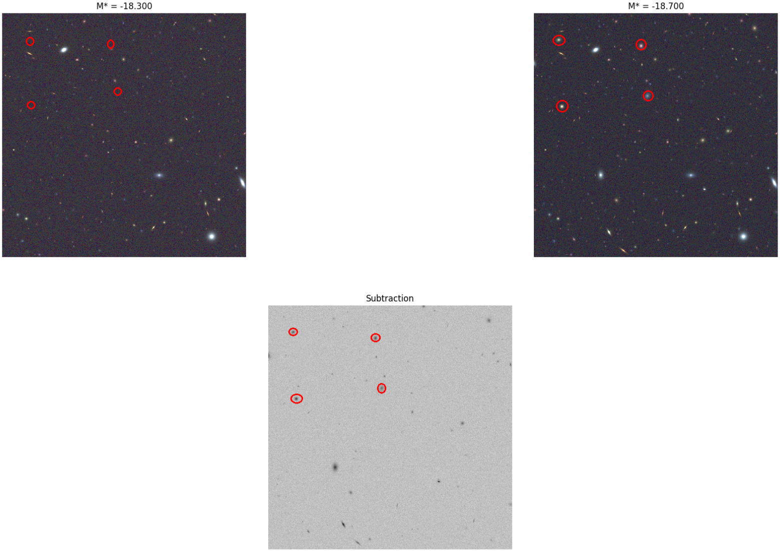
- Let X be a random variable with density f depending on parameters Θ, the Fisher information carried by X about the parameters Θ is defined by:

$$F(\Theta) = \mathbb{E}\left[\left(\frac{\partial}{\partial\Theta}\log f(X,\Theta)\right)^2 \middle| \Theta\right] = \int \left(\frac{\partial}{\partial\Theta}\log f(x,\Theta)\right)^2 f(x,\Theta)dx \tag{5}$$

- Fisher information = quantity of information of a set of images depending on a set of parameters α, M^*, ϕ^*, M^*_{evol} and ϕ^*_{evol}.
- Depends on the **covariance matrix** and on the **derivative of the mean** for the parameters.
- Résumé: **statistical impact of a slight perturbation of the parameters on the set of images**.

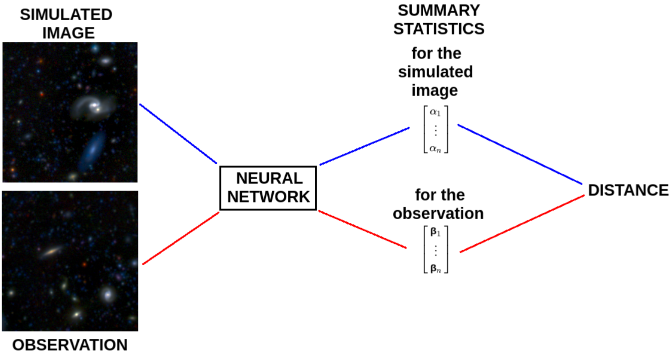
LOSS FUNCTION: $\Lambda = -\ln|\mathbf{F}(\Theta)|$

Residual image



Comparison of simulated and observed images

Charnock et al. 2018: "Automatic physical inference with information maximising neural networks".

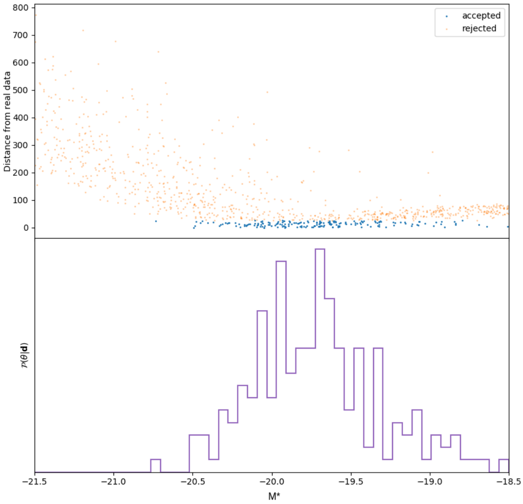


$$\text{distance} = \sqrt{(\alpha - \beta)^T F (\alpha - \beta)} \tag{6}$$

where F is the Fisher information matrix obtained by the trained network.

First results

Real data obtained for $M^* = -19.94$.



Conclusion

- Neural network algorithm upgraded by Tom Charnock.
- Consistent simulated images.
- Next results to come.

THANK YOU FOR YOUR ATTENTION